

DIAS—Earth Environment Data Integration and Analysis System

BY EIJI IKOMA AND MASARU KITSUREGAWA

OUR GROUP HAS been developing and operating a platform to acquire, archive, and manage various data related to the Earth's environment to make it available to researchers across a wide range of fields.

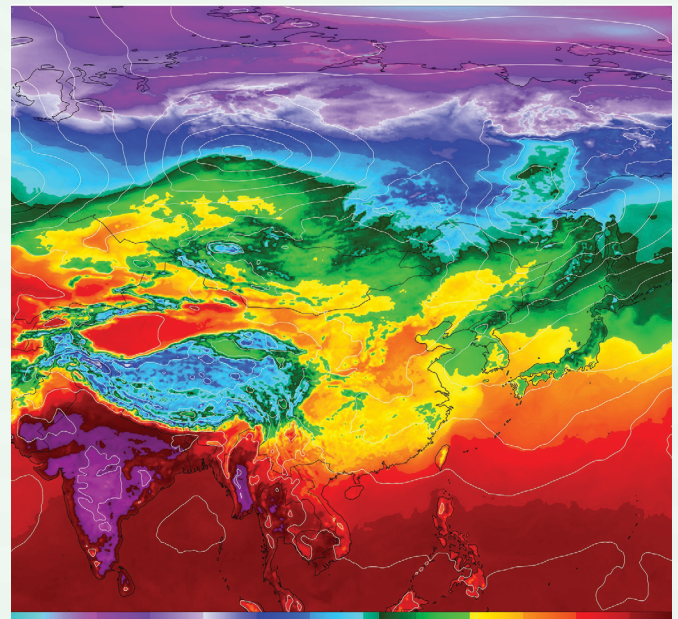
Development of this system began in the 1980s to receive, archive, and distribute Asian satellite image data. Currently, the system covers a variety of data including weather, climate change, disaster prevention, biodiversity, health, and agriculture. Today, the Data Integration and Analysis System (DIAS^a) is a large-scale analysis platform with huge storage and more than 10,000 registered users (half of them in Japan and the other half primarily in Asia). As shown in the accompanying figure, users can easily use data collected by the common collection API through the common use API, and they can operate services at the application layer.

Archived real-time data

a <https://diasjp.net/en/>

includes observation data from automatic observation equipment, weather radars, live cameras, and forecast data generated on occasion. In many cases, once an acquisition fails, it cannot be completed. Therefore, we developed a robust tool to ensure data acquisition within a specified timeframe being used for river observation data, live camera data, and radar data. Social sensing data is another type of real-time data, for which we use a common platform developed for biodiversity research and disaster information collection applications. Non-real-time data includes observational, experimental, and simulation data generated by large-scale forecasting experiments. For efficient data uploading, we provide our own quality control and metadata input tools, which have achieved rapid data release in many international research projects, significantly shortening the time from data acquisition to data release.

The data download system, which is the most



frequently used function, has been improved to add a function to extract only the necessary elements in consideration of the recent increase in data volume, and an API is also provided for real-time users. In addition, we offer an analysis environment (virtual server environment) that can directly access real-time data and very large amounts of data that is difficult to download. It allows users to login to the DIAS system and directly develop and operate services on it. This service not only reduces the enormous download time but also enables real-time analysis in an environment where data is acquired in real time.

Currently, similar data infrastructures exist, such as Copernicus in Europe and Open Data Cube in Australia. They may lack the ability

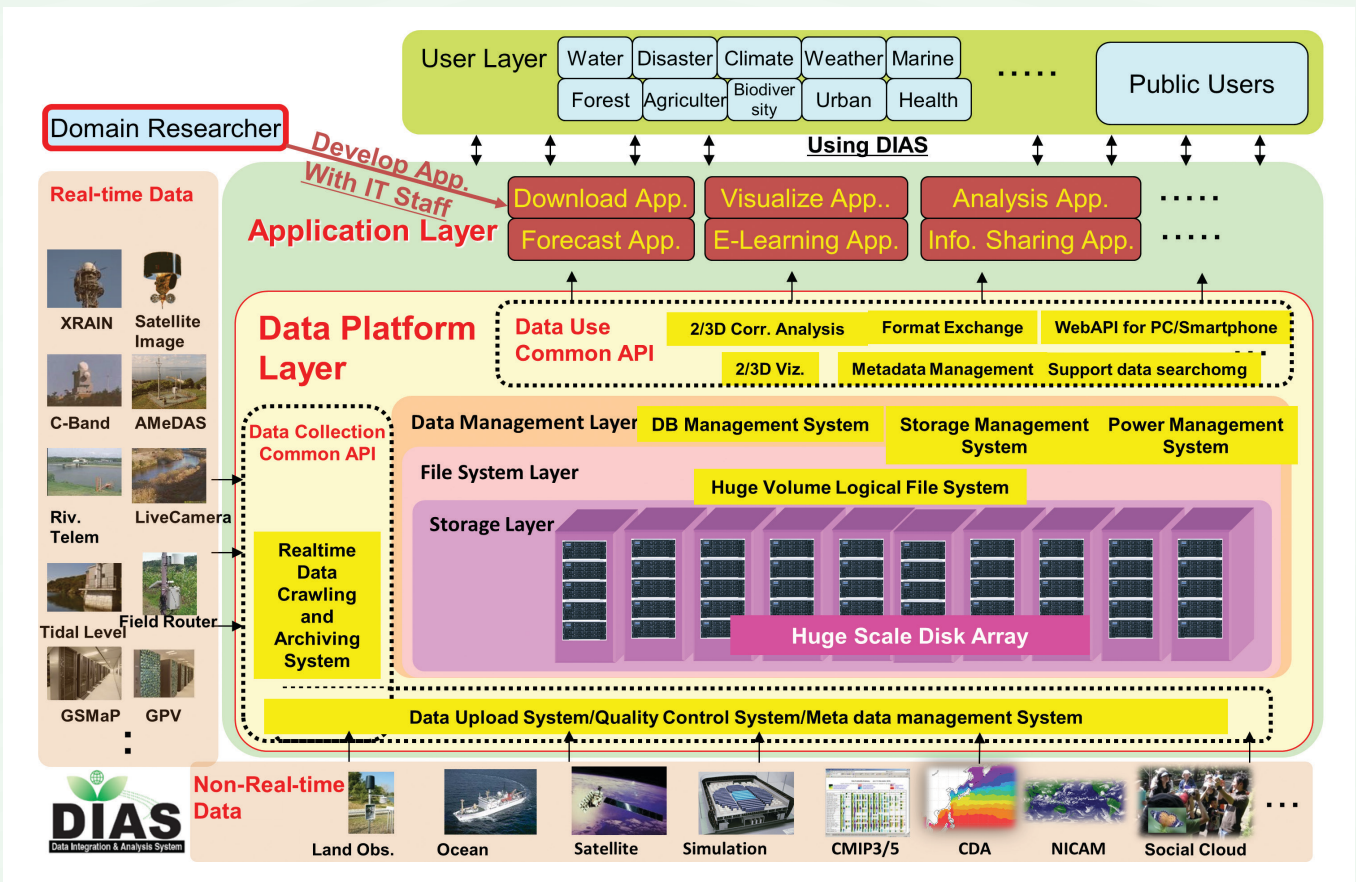
to integrate different types of data or be specialized systems for specific types of data. Neither of the two systems can provide integrated data use and services as flexibly as DIAS.

Several applications that use this infrastructure are currently in operation. For example, a service that predicts river flooding and inundation by integrating and analyzing various types of real-time data archived by the DIAS and local observation data is in operation in the Philippines^b and Sri Lanka,^c where floods have occurred frequently in recent years. It is being expanded to Vietnam, Indonesia, Myanmar, Iran, and

b <https://www.mdpi.com/2073-4441/13/9/1218>

c <https://www.mdpi.com/2073-4441/14/6/978>

DIAS is a large-scale analysis platform with huge storage and more than 10,000 registered users.



DIAS system structure.

the Niger-Volta River basin in Africa, where similar damage has occurred. In addition to issuing forecast information, the system is also equipped with an e-learning system for policy-makers of local institutions to learn how to operate the system, thereby contributing to human resource development in each country. In addition, the database on butterflies developed using DIAS's citizen-participatory data collection platform has been used to contribute to research on biodiversity in Japan.^d


Because this platform is species-independent, the system is now being used for many other organisms, such as fish, shellfish, dragonflies, and storks, contribut-

ing significantly to the fields where databases have not been adequately developed. S-uIPS,^e a real-time urban inundation forecasting application that utilizes minute-by-minute rainfall data, rainfall forecast data, and vast amounts of city block data, also uses this platform. It is now available for Tokyo^f in response to urban disasters associated with heavy rainfall, which are particularly severe in the Asian region. This service had used a program developed by a researcher in social infrastructure engineering to calculate road flooding depths, but it took approximately 10 hours to make a 30-minute prediction. Our develop-

e https://www.jstage.jst.go.jp/article/jscejhe/75/2/75_I_1315/_article/-char/ja/
f <https://diasjp.net/service/s-uips/>


ment team, consisting of experts in information engineering, assisted in re-writing the program to take advantage of GPUs, which resulted in an improvement of efficiency by 60 times, enabling real-time operation. This is a good example of cooperation between non-IT and IT experts and is a result of interdisciplinary collaboration. Because this service does not use region-specific parameters, it can be easily deployed in other Asian cities with problems like inland flooding, as long as the city's infrastructure information is available.

Thus, we not only provide and expand the data on DIAS but also collaborate with researchers in various fields to provide an efficient usage environment and cooperate to expand the already obtained solutions to more countries. In the future, we

plan to respond to requests from data providers and support the analysis of user trends by employing access logs and other data. We will continue to expand and improve the functions of DIAS to support the research community in Asia that uses Earth environmental data. Particularly, we are analyzing the interrelationships among the data fields used, the relationship between the occurrence of sudden social events and high access data, and the design of efficient interfaces based on typical access patterns. 

Eiji Ikoma is Project Associate Professor for the Earth Observation Data Integration and Fusion Research Initiative (EDITORIA) at the University of Tokyo.

Masaru Kitsuregawa is Project Professor for the Earth Observation Data Integration and Fusion Research Initiative (EDITORIA) at the University of Tokyo.

 This work is licensed under a <https://creativecommons.org/licenses/by/4.0/>